HW

* I searched Kaggle for a csv with the cost of avocados from November 15th, 2015 to February 11th, 2018.
* I calculate statistics for the average price of avocados.
* My data is quantitative, discrete and ratio. Quantitative because my data is numerical. Discrete because cost can only take certain values (prices with two numbers after the decimal points max), and ratio because a true zero exists. The price of avocado can be 0.00, if the avocado was simply free.

Descriptive Stats

|  |  |
| --- | --- |
| *Stats for avocado prices* | |
| Mean | 1.36492061 |
| Standard Error | 0.00207505 |
| Median | 1.18219 |
| Mode | 1.33 |
| Standard Deviation | 0.28031572 |
| Sample Variance | 0.0785769 |
| Kurtosis | -0.7446997 |
| Skewness | 0.85050102 |
| Range | 0.8999 |
| Minimum | 1.1 |
| Maximum | 1.9999 |
| Sum | 24908.4362 |
| Count | 18249 |

|  |  |
| --- | --- |
| Manual stats for avocado prices | |
| Significance level | 0.05 |
| Mean | 1.36 |
| Median | 1.18219 |
| STDEV | 0.28031572 |
| Mode | 1.15 |
| Range | 0.8999 |
| Sum | 24908.4362 |
| Sample size | 18249 |
| Margin of error | 0.00406702 |
| Confidence interval | 1.36 +0.00406702 |

A picture containing chart

Description automatically generated

Descriptive Statistics Description

The size of my sample is relatively large- 18249. The median is lower than the mean, meaning that most of the values in the data set are lower than the average. The standard deviation is smaller, meaning that the data does not vary widely from the mean.  My data is skewed to the right. There are no outliers in the data